Available online at www.sciencedirect.com

SCIENCE DIRECT

International Journal of Industrial Ergonomics 34 (2004) 407–419

International Journal of Industrial Ergonomics

www.elsevier.com/locate/ergon

Measurement of human trust in a hybrid inspection system based on signal detection theory measures

Xiaochun Jiang[a,*], Mohammad T. Khasawneh[b], Reena Master[c], Shannon R. Bowling[c], Anand K. Gramopadhye[c], Brian J. Melloy[c], Larry Grimes[d]

[a] *Department of Industrial and Systems Engineering, North Carolina A&T State University, 1601 E. Market Street, Greensboro, NC 27411, USA*

[b] *Department of Systems Science and Industrial Engineering, State University of New York at Binghamton, Binghamton, NY 13902, USA*

[c] *Department of Industrial Engineering, Clemson University, Clemson, SC 29634-0920, USA*

[d] *Department of Experimental Statistics, Clemson University, Clemson, SC 29634-0920, USA*

Received 17 June 2003; accepted 7 May 2004

Available online 26 June 2004

Abstract

Human trust plays an important role in influencing operator's strategies toward the use of automated systems. Therefore, a study was conducted to measure the effect of human trust in a hybrid inspection system given different types of errors (i.e., false alarms and misses). The study also looked at which of the four dimensions of trust (competence, predictability, reliability and faith) were the best predictors of overall trust. Results from the study reveal that trust is sensitive to the type of errors made by a system and suggest that subjective ratings of trust and the properties of the system can be used to predict the allocation of functions in hybrid inspection systems.

Relevance to industry

The study conducted here is applicable to inspection tasks in manufacturing and service industries. The results obtained indicate that subjective ratings of operators' trust can be used as a basis for predicting and optimizing operator's allocation behavior and system performance. Furthermore, designers can use these results to help decide which functions to allocate to the human or to the system based on previous experiences and interaction with the system.

Keywords: Trust; Function allocation; Hybrid inspection

1. Introduction

Customer awareness regarding product quality and increased incidences of product liability litigation has increased the importance of the inspection process in manufacturing industries

*Corresponding author. Tel.: +1-336-334-7780; fax: +1-336-334-7729.

E-mail addresses: xjiang@ncat.edu (X. Jiang), agramop@ces.clemson.edu (A.K. Gramopadhye).

0169-8141/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ergon.2004.05.003

(Thapa et al., 1996). To remain competitive, manufacturers can accept only extremely low defect rates, often measured in parts per million. This situation requires almost perfect inspection performance in the search for nonconformities in a product, and the two functions central to inspection, visual search and decision making, have been shown to be the primary determinants of inspection performance (Sinclair, 1984; Drury, 1992). However, while the need for error-free detection is important, the inspection process tends to be less than 100% reliable, especially when human inspectors are used (Drury, 1992). As a result, many companies are moving toward automated inspection systems (Thapa et al., 1996). Even though these automated systems have well-documented advantages, they cannot surpass the innate ability of humans to recognize patterns, make rational decisions, and quickly adapt to new situations (Hou et al., 1993; Gramopadhye et al., 1997). It is possible, though, that superior performance could be achieved with a hybrid system, one in which automation complements human strengths. One challenge in designing such an inspection system is determining how best to allocate functions between humans and machines, an especially critical issue since research has shown that proper function allocation has the potential to lead to improvement in inspection performance (Bullinger and Salvendy, 1987; Sharit and Elhence, 1987; Morris et al., 1988; Sinclair, 1993; Jiang et al., 2004). In response to this need, Hou et al. (1993) proposed the seven alternate hybrid inspection systems listed in Table 1 with Alternative Seven, the most complicated and flexible, chosen for use in the current study.

To measure system performance, traditional measures of speed, accuracy were often used (Hou et al., 1993). In addition to these, several studies of supervisory control have confirmed that trust in automation, a subjective measure, is a key component in determining how to allocate functions effectively (Muir and Moray, 1996; Lee and Moray, 1992; Sheridan, 1980; Sheridan et al., 1983a,b; Sheridan and Hennessy, 1984). Research from both social science and engineering viewpoints suggests that trust is a multidimensional concept, reflecting a set of interrelated perceptions such as the reliability or the predictability of an entity, the behavior of the human involved, the characteristics of the machine, and the interactions between the operator and the system (Bisantz et al., 2000; Parasuraman, 2000; Moray and Rodriguez, 2000).

The first step in gaining a better understanding of how the characteristics of automation ultimately affect human behavior and, as a result, inspection performance, is to identify the factors influencing trust. In the absence of human/machine models of how trust evolves, models between people have been used as a basis. Muir (1994) was one of the first to develop such a model using this approach. He attempted to capture the multi-dimensional construct of trust by incorporating dimensions that evolved from two models of trust between humans: Barber's (1983) model and Rempel et al. (1985) model.

Table 1
Allocation alternatives in hybrid inspection task (Hou et al., 1993)

Alternative	Search	Decision-making	System mode
1	Human	Human	Human
2	Computer	Computer	Computer
3	Human	Computer	Hybrid
4	Computer	Human	Hybrid
5	Human	Human + Computer	Hybrid
6	Computer	Human + Computer	Hybrid
7	Human + Computer	Human	Hybrid
8	Human + Computer	Computer	Hybrid
9	Human + Computer	Human + Computer	Hybrid

Barber (1983) defined trust in terms of a taxonomy incorporating three expectations: persistence, technical competence, and fiduciary responsibility, with only the first two applicable to machine–human relationships. For him, persistence, the foundation for trust, was established by an expectation of constancy. In adapting this construct to automation, Muir (1994) proposed that constancy is what allows humans to predict the future state of a system while the expectation of technical competence, what humans probably most easily understand as the meaning of trust in machines, is the proper execution of a specific task.

In addition to these two dimensions proposed by Barber (1983), Muir (1994) incorporated three from Rempel et al. (1985), whose model consists of a hierarchy with trust at any one stage based on the outcome of the earlier one. Using Rempel et al.'s three components of predictability, dependability and faith, Muir proposed that, similar to the development of trust between people, humans begin to develop trust in a machine by determining its predictability, i.e., by evaluating the consistency and desirability of its repeated behavior over a given period of time. As the relationship progresses, this predictability leads to dependability, the second stage in the hierarchy. The development of faith, which according to Rempel et al. requires the taking account past experiences based on predictability and dependability, occurs over time, suggesting humans develop faith only after working with a machine for significant amount of time.

By integrating Barber's and Rempel et al.'s model to include the dimensions of persistence, technical competence, predictability, dependability and faith, Muir (1994) was able to establish a more comprehensive model of the development of human trust in automation. The model was later reduced to the four dimensions of trust found to be applicable to hybrid inspection systems (Master et al., 2000).

In addition to understanding the relationship between trust and its dimensions, it is equally important to be able to measure these effectively. Previous research has investigated various methods for measuring trust, primarily utilizing rating scales in questionnaires (Muir and Moray, 1996; Singh et al., 1993). However, the applicability of the results obtained from these early questionnaires is limited because some measured trust in a particular person while others were general in nature (Bisantz et al., 2000). Furthermore, they were based on theoretical notions of trust rather than empirical analysis. In response to these limitations, Jian et al. (2000) developed an empirically determined scale for measuring trust in automated systems to understand better the similarities and differences in the concepts of trust. While an improvement, it was based on the measurement of trust in any general automated system and, as a result, is not completely applicable to hybrid inspection systems. Consequently, a new questionnaire was developed to determine the effects of the level of trust an operator has in hybrid inspection systems for administration for the current study (Master et al., 2000). This questionnaire incorporated the four dimensions of trust—competence, predictability, reliability and faith—derived from the multidimensional construct developed by Muir (Master et al., 2000) to determine the best predictors of trust, the first objective of this study, with the questionnaire developed by Jian et al. (2000) being used to validate the questions common to both questionnaires.

The main objective of this paper was to measure the effects of human trust in a hybrid inspection system. As Muir and Moray (1996) discovered in their study of nine pumps, operator trust is affected by accuracy, i.e. the type and number of errors made by the system. Therefore, using the signal detection theory classification of systems as conservative, risky, or neutral (Swets, 1964; Green and Swets, 1966), this study measured the effects of misses, and false alarms on human trust of automation in a hybrid inspection system. The study also looked at which of the four dimensions of trust (competence, predictability, reliability and faith) were the best predictors of overall trust.

Section 2 of this paper describes the methodology used to achieve the two objectives of this study, detailing the participants, the stimulus material, the inspection task, the experimental design, the procedure, and the data collection. Section 3 presents the results concerning the effect of system error (false alarms and misses) on human trust of automation as well as the dimensions and predictors of trust, while Section 4 discusses these findings and offers conclusions.

2. Methodology

2.1. Subjects

The subjects were 6 students, both graduate and undergraduate, enrolled at Clemson University between the ages of 18 and 28. Students can be used as subjects in lieu of inspectors because as Gallwey and Drury (1986) have shown, minimal differences exist between inspectors and student subjects on simulated tasks. The subjects were screened for 20/20 vision, corrected if necessary, and paid \$5.00/h for their time.

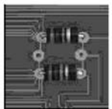
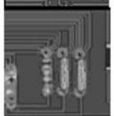
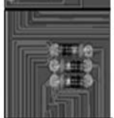
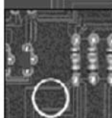
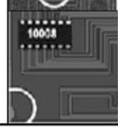
2.2. Stimulus material

The task was a simulated visual inspection of a printed circuit board implemented on a Pentium III computer with a 19″ high-resolution (1024 × 768) monitor. The input devices were a Microsoft standard keyboard and a Microsoft one-button mouse. The task consisted of inspecting simulated PCB images developed using Adobe PhotoShop 5.5 for six categories of defects seen in Table 2—missing components, wrong components, inverted components and misaligned components, trace defects and board defects. Four categories of defects could occur on any of these four individual components: resistors, capacitors, transistors and integrated circuit.

2.3. Inspection systems

Human+computer search/human+computer decision-making hybrid inspection system: In this

Table 2
Illustrations of defects

	Defect category	Defect type		Sample defect
1	Missing component	Missing	Resistor[a] Capacitor Transistor IC	
2	Wrong component	Wrong	Resistor[a] Capacitor IC/Board number Transistor	
3	Misaligned components	Inverted	Resistor[a] Capacitor IC Transistor	
4	Polarity errors	Slightly/ Significantly Misaligned	Resistor Capacitor[a] IC Transistor	
5	Trace defects	Open[a]/Short Slight/ significant Soldering Joints	Copper Overlay	
6	Edge defect	Slight/ Significantly	Board Defect[a]	

[a]Corresponds to illustration

system, both computers and humans searched for defects and made the decision on the board with the human having the final decision about whether to accept or override the computer search or decision-making (Jiang et al., 2002, 2003). During visual search, PCB boards containing 1, 2, 3, or no defects were presented to the subjects, whose task was to locate all potential defects and name them. After locating the defect, they clicked the mouse on it and chose its name from a dropdown box listing all possible defects. At the same time, the computer performed the same search task. However, subjects could override the computer if they did not agree with its search results. Then, the computer made its conformance decision and the subjects made their final conformance decisions, either agreeing or disagreeing with the computer. Once the board was classified, the image of the next board would be presented to the subjects. Each inspection task consisted of 48 randomly ordered PCB boards—12 of each zero-defect, single-defect, two-defect, and three-defect boards. Table 3 provides detailed information about the boards while Fig. 1 shows a typical decision-making response by the computer and the human inspector's decision to override this decision.

2.4. Experimental design

This study used a single factor (response bias) within subject design. The three levels of the response bias were conservative (high false alarms/low misses), neutral (equal false alarms and misses), and risky (high misses/low false alarms). Table 4 shows the layout of the design. Two Latin squares with different orders were used to cancel out the order effects (see Table 5). All treatments were randomly assigned to the three Latin letters.

2.5. Procedure

The study took place over a 7-day period. Day one was devoted to training the subjects and during the next 6 days, data were collected on the criterion tasks. A more detailed explanation of the activities conducted on each day can be seen below.

Table 3
Details of selected PCB boards

Board type	Number of boards	Defect type
Conforming boards	12	None
Single-defect boards	2	Missing component
	2	Wrong component
	2	Inverted component
	2	Misaligned component
	2	Board defect
	1	Copper overlay
	1	Soldering joint
Two-defect boards	2	Missing component and wrong component
	2	Missing component and misaligned component
	2	Wrong component and misaligned component
	2	Missing component and copper overlay
	2	Misaligned component and board defect
	2	Inverted component and wrong component
Three-defect boards	3	2 wrong components and 1 misaligned component
	3	1 each missing component, wrong component and misaligned component
	3	1 each board defect component, copper overlay component, and misaligned component
	3	1 each wrong component, inverted component, and wrong component

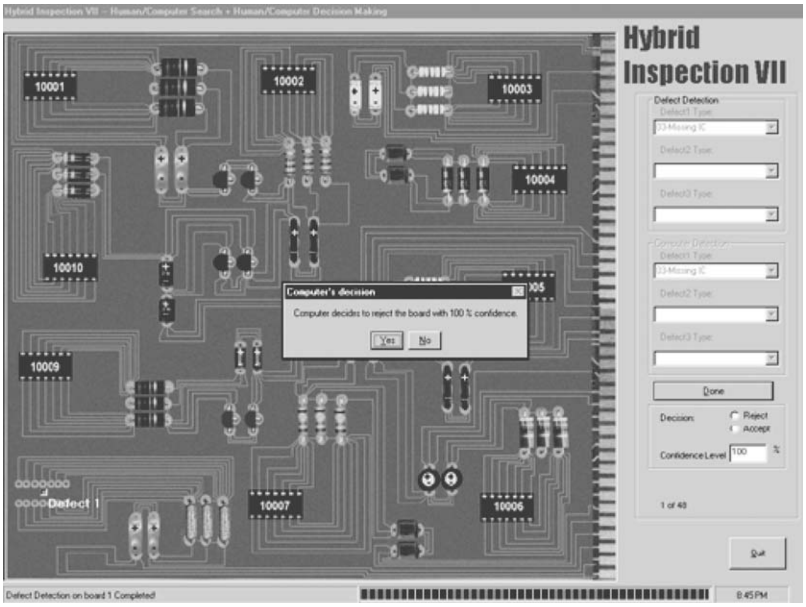


Fig. 1. Screenshot from Hybrid Inspection System VII.

Table 4
Layout of the design

Hybrid inspection system		
Conservative	Neutral	Risky
	Subject 1 | Subject 6	

Table 5
Latin square design

		Subjects					
		1	2	3	4	5	6
Sequence of experimental conditions	1	C	B	A			
	2	A	C	B			
	3	B	A	C			
	4				B	C	A
	5				C	A	B
	6				A	B	C

On Day one, each subject was required to complete a consent form and a demographics questionnaire. Following this step, instructions were read to the subjects to ensure their understanding of the experiment. Next, all were trained and given three separate tests before beginning the experiment.

A typical training session proceeded as follows:

1. *Initial overview*: Initially, the subjects were introduced to basic PCB inspection terminology and familiarized with the computer program. Following this step, subjects were quizzed on their knowledge of the operation of the software, and correct answers were supplied for incorrect responses.
2. *Defect training*: The subjects were initially trained to recognize different types of defects by being shown instances of each, including name and probable locations. Then, training was provided on the guidelines used to classify the PCB board as conforming or nonconforming.

After completion of defect training, the subjects underwent training sessions on defect matching, single defect inspection and multiple-defect training. Following each session, the subjects were administered a test. Only those subjects who secured a minimum score were allowed to proceed to the next step.

3. *Defect matching*: PCBs with a marked single defect were displayed on the screen, and subjects classified it by choosing the correct name from a dropdown box. They were provided with immediate feedback about the correctness of their responses.
4. *Single-defect training*: PCBs with a single defect were displayed on the screen, and subjects located and then classified it by choosing the name from a dropdown box. They were provided with immediate feedback on their search performance using speed and accuracy measures.
5. *Multiple-defect training*: PCBs with 1, 2 or 3 defects were displayed on the screen, and subjects first visually searched for and then classified them. They were provided with immediate feedback on their performance using speed and accuracy measures.

On Day two, to develop a baseline of a subject's trust of the system, each was administered a criterion task with 24 PCB boards to inspect using a perfect hybrid inspection system, i.e., a system did not make any errors.

On Day three, subjects were required to fill out a trust questionnaires at three different times. At the first stage, they were required to complete a generic trust questionnaire (Jian et al., 2000) and a trust questionnaire developed by Master et al. (2000) to measure initial trust before they began the experiment. Following this step, they were administered one trial block of 48 PCB boards based on one of the three treatments. In stage two, the subjects were asked to complete the same two trust questionnaires after they finished 24 boards of each trial block. In stage three, they were required to complete another set of the same questionnaires after they finished the inspection task. From Day four to Day seven, the subjects followed the same procedure and were assigned the other two experimental conditions. On completion of the study, each subject was debriefed.

3. Results

The results from the hybrid inspection trust questionnaire were first compared with those from the generic questionnaire using correlation analysis. Then, they were analyzed using the mixed model analysis of covariance to analyze the effect of the system response bias on trust measurement at stage two, after the subjects finished 24 boards, and at stage three, on completion of the inspection task. Finally, stepwise regression analysis was used to select the best predictors of overall trust at each of three stages as well as the change in the trust for the three inspection systems.

3.1. Correlation analysis

Since the generic questionnaire developed by Jian et al. (2000) has been validated, a correlation analysis was conducted on the two trust questionnaires for all three stages to validate the hybrid inspection trust questionnaire as shown in Table 6.

As this analysis shows, the hybrid inspection questionnaire used in this study has a very high correlation with the generic questionnaire, indicating that the hybrid inspection questionnaire has been validated.

3.2. Analysis of covariance for the trust

3.2.1. Stage two-after 24 PCB boards

Overall: As illustrated in Fig. 2, the analysis of covariance of overall trust indicated a significant treatment effect ($F(2,27) = 7.71, p < 0.01$). The results of Student–Neuman–Keuls procedure are summarized in Table 7. Treatments underlined by a common line do not differ from one another while treatments not underlined by a common line do. As shown in Table 7, the overall trust of the risky system is significantly lower than the neutral system and the conservative system while the latter two are not significantly different.

Table 6
Correlation analysis for the two questionnaires

	Correlation (r^2)	p
Stage one	0.97	<0.001
Stage two	0.975	<0.0001
Stage three	0.98	<0.001

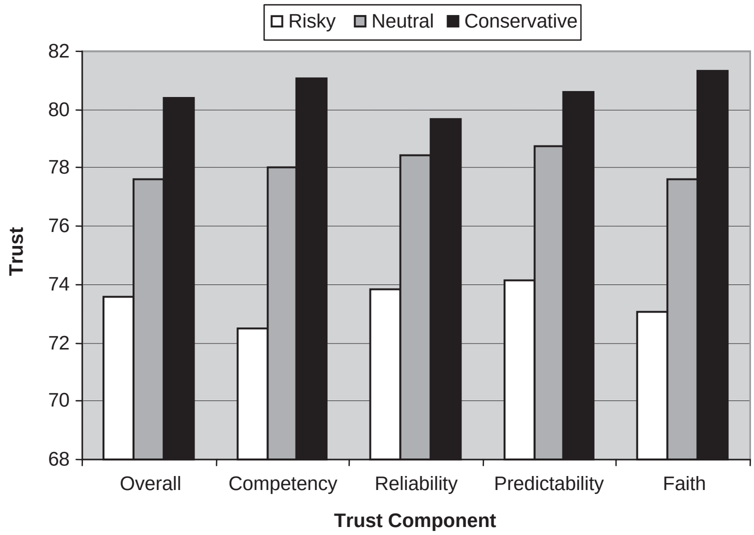


Fig. 2. Trust at stage two.

Table 7
Results of Student–Neuman–Keuls procedure for three inspection systems

Stage	Component	Tests on all ordered pairs of means		
2	Overall trust	Risky	Neutral	Conservative
	Competency	Risky		
			Neutral	Conservative
	Reliability	Risky		
			Neutral	Conservative
	Predictability	Risky	Neutral	Conservative
	Faith	Risky	Neutral	Conservative
3	Overall trust	Risky	Neutral	Conservative
	Competency	Risky		
			Neutral	Conservative
	Reliability	Risky		
			Neutral	Conservative
	Predictability	Risky		
			Neutral	Conservative
	Faith	Risky		
			Neutral	Conservative
		Risky		
			Neutral	Conservative

Competency: As illustrated in Fig. 2, the analysis of covariance of competency indicated a significant treatment effect ($F(2,27) = 7.16, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that competency of the risky system is significantly lower than the neutral system and the conservative system while the latter two are not significantly different.

Reliability: As illustrated in Fig. 2, the analysis of covariance of reliability indicated a significant treatment effect ($F(2,27) = 7.16, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that reliability of the risky system is significantly lower than the neutral system and the conservative system while the latter two are not significantly different.

Predictability: As illustrated in Fig. 2, the analysis of covariance of predictability indicated a significant treatment effect ($F(2,27) = 5.10, p < 0.05$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that predictability of the risky system is significantly lower than the conservative system while it is not significantly different from the neutral system.

Faith: As illustrated in Fig. 2, the analysis of covariance of faith indicated a significant treatment effect ($F(2,27) = 8.64, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that faith of the risky system is significantly lower than the conservative system while it is not significantly different from the neutral system.

3.2.2. Stage three-after 48 PCB boards

Overall: As illustrated in Fig. 3, the analysis of covariance of overall trust indicated a significant treatment effect ($F(2,27) = 8.87, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that overall trust of the conservative system is significantly higher than the neutral system and the risky system while the latter two are not significantly different.

Competency: As illustrated in Fig. 3, the analysis of covariance of competency indicated a significant treatment effect ($F(2,27) = 8.48, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that competency of the risky system is significantly lower than the neutral system and the risky system while the latter two are not significantly different.

Reliability: As illustrated in Fig. 3, the analysis of covariance of overall trust indicated a significant treatment effect ($F(2,27) = 7.54, p < 0.01$).

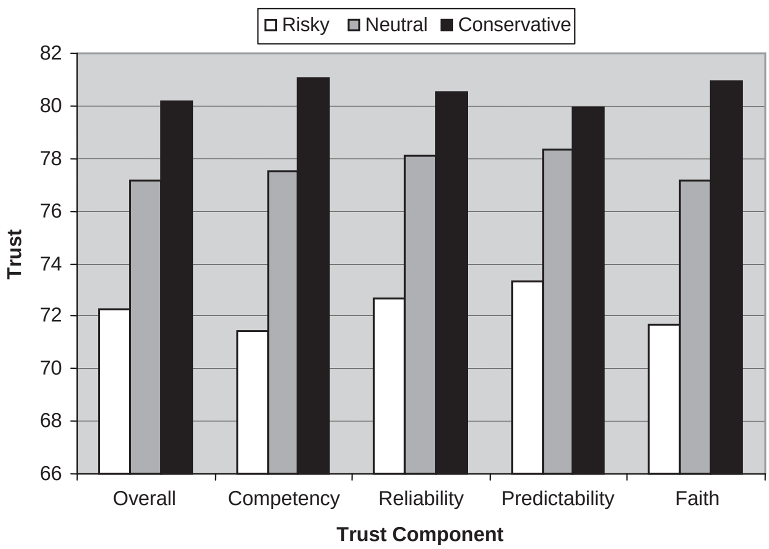


Fig. 3. Trust at stage three.

The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that reliability of the risky system is significantly lower than the neutral system and the risky system while the latter two are not significantly different.

Predictability: As illustrated in Fig. 3, the analysis of covariance of predictability indicated a significant treatment effect ($F(2,27) = 6.48, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that predictability of the risky system is significantly lower than the neutral system and the risky system while the latter two are not significantly different.

Faith: As illustrated in Fig. 3, the analysis of covariance of faith indicated a significant treatment effect ($F(2,27) = 9.09, p < 0.01$). The results of Student–Neuman–Keuls procedure summarized in Table 7 showed that faith of the risky system is significantly lower than that of the neutral system and that of the risky system while the latter two are not significantly different.

3.3. Stepwise regression model analysis

The stepwise regression procedure was used to analyze trust model and to find out the best

Table 8
Stepwise regression analysis for the risky system

	Initial trust (stage one)	Middle trust (stage two)	Final trust (stage three)	Change in trust
R^2	0.9880	0.9994	0.9843	0.9661
$C(P)$	1.2964	1.5793	1.4765	1.1823
Best predictors	Faith	Reliability and predictability	Faith and reliability	Faith and reliability
Model(T)				
Competency				
Reliability		0.9395	−0.0829	−0.1055
Predictability		0.0567		
Faith	0.9804		1.073	1.3480
Intercept	2.002	0.1966	1.9487	1.165
F	330.07	2602.47	94.21	42.74
p	<0.0001	<0.0001	<0.01	<0.01

Table 9
Stepwise regression analysis for the neutral system

	Initial trust (stage one)	Middle trust (stage two)	Final trust (stage three)	Change in trust
R^2	0.9685	0.9994	0.9843	0.9661
$C(P)$	184.5293	1.5793	1.4765	1.1823
Best predictors	Competence	Reliability and predictability	Faith and reliability	Faith and reliability
Model(T)				
Competency	1.3984			
Reliability		0.4441	3.6944	
Predictability			0.5066	
Faith		0.7475	−2.4349	0.6185
Intercept	−32.35	−12.4511	18.008	−3.0809
F	123.14	137.68	60.35	242.62
P	<0.01	<0.01	<0.05	<0.0001

Table 10
Stepwise regression analysis for the conservative system

	Initial trust (stage one)	Middle trust (stage two)	Final trust (stage three)	Change in trust
R^2	0.9803	0.9994	0.9843	0.9661
$C(P)$	1.2228	1.5793	1.4765	1.1823
Best predictors	Faith	Reliability	Faith and reliability	Faith and competency
Model(T)				
Competency	1.3984			−0.6896
Reliability		1.2795	1.4508	
Predictability				
Faith	1.1256		0.3052	1.5489
Intercept	11.4861	18.233	−54.6064	1.165
F	198.91	21.24	164.23	55.12
p	<0.001	<0.05	<0.01	<0.01

predictors at each of three stages as well as the change in the trust for the three inspection systems. Results are reported in Tables 8–10.

4. Discussion and conclusions

The most salient finding of this research is that measuring trust in a hybrid inspection environment is feasible and that the trust questionnaire is sensitive to different inspection systems.

At stage two, it was found that the overall trust in the risky system was significantly lower than that of the other two systems while no significant difference was found between those two systems. This may have been a result of the percentage of the types of errors made by each system; for example, the risky system had a high percentage of false alarms and a low percentage of misses. Since false alarms made by the system are commissive errors, which the subjects cannot overlook if they are paying attention, whereas misses are omissive errors which the subjects are likely to miss if they are not paying attention (Sanders and McCormick, 1993), people's overall trust of a computer with a risky bias may be affected more than it would be for computers with a conservative or neutral bias. Furthermore, since a neutral system makes only slightly more false alarms than a conservative system, it may not be able to draw an operator's attention and, therefore, no significant difference is reflected in the trust of the systems. The same pattern was revealed from the analysis of covariance at stage two for two of the four dimensions of the trust: competency, which is "the extent to which the system performs the task effectively" (Master et al., 2000), and reliability, which is "the extent to which the system is free of errors" (Master et al., 2000).

On the other hand, the analyses for the other two dimensions, predictability and faith, gave a different pattern. Although it remained true that the trust of the neutral system and the conservative system were statistically the same, no significant difference was found between the trust of a risky system and that of a neutral system. Probably because predictability and faith are more sensitive to system errors, a slight change in the error rate may have some effect on the trust in the system. Furthermore, at stage two, subjects probably did not have a clear sense of the slight difference between a risky system and a neutral system or a neutral system and a conservative system. As a result, although predictability and faith in a risky system were significantly lower than those in a conservative system, there is no significant difference between them in both a risky system and a neutral system, or in a conservative system and a neutral system.

At stage three, it was found that the overall trust as well as the four dimensions of the trust in the risky system was significantly different in the other two systems while no significant difference was found among them in the latter two systems. The difference between stages two and three is that predictability and faith also follow the same pattern as the other dimensions as well as overall trust in stage three. One possible reason is that subjects became more aware of the difference in the risky and the neutral systems at stage three and, therefore, rated them more differently than they did at stage two. As a result, trust of the risky system is significantly different from that of the neutral system.

Results from the stepwise regression analysis conducted at three stages for the three inspection systems revealed that a linear regression model could be used to predict operator's trust of a hybrid inspection system. For example, faith and reliability are the best predictors for all three systems at stage three.

When looking at the subject's ratings of overall trust, it was seen that there was a high degree of variability between each subject's rating of trust. This shows that trust is subjectively based on the characteristics of the operator. To understand better how trust is affected by system errors, individual differences resulting from the different types of human behavior need to be taken into account.

Another reason for the variance in trust between subjects was probably because none of the subjects had been exposed to this type of system, even though each subject was trained to use it prior to starting the study. Since subjects probably based their initial trust on past experiences encountered with similar systems, it can be hypothesized that operators need to be experienced users of the system in order to be able to elicit without much variability.

These results show that operators' subjective ratings of trust can provide insight into their relationship with the system and can be used as a basis for predicting and optimizing allocation behavior and system performance. They are helpful for both designers and researchers. Designers of automation can use them to help decide which functions to allocate to the human or to the system based on previous experience and interaction with the system. Researchers can use them to perform the field studies needed to validate laboratory results to understand further human trust in automation.

References

Barber, B., 1983. The Logics and Limits of Trust. Rutgers University Press, New Brunswick, New Jersey.

Bisantz, A.M., Llinas, J., Seong, Y., Finger, R., Jian, J., 2000. Empirical investigations of trust-related system vulnerabilities in aided, adversarial decision-making. Department of Industrial Engineering, State University of New York, Buffalo.

Bullinger, H.J., Salvendy, G., 1987. Human aspects of robotic systems. In: Salvendy, G. (Ed.), Handbook of Human Factors. Wiley, New York.

Drury, C.G., 1992. Inspection performance. In: Salvendy, G. (Ed.), Handbook of Industrial Engineering, 2nd Edition. Wiley, New York.

Gallwey, T.J., Drury, C.G., 1986. Task complexity in visual inspection. Human Factors 28 (5), 595–606.

Gramopadhye, A.K., Drury, C.G., Prabhu, P.V., 1997. Training for visual inspection. Human Factors and Ergonomics in Manufacturing 7 (3), 171–196.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Hou, T., Lin, L., Drury, C.G., 1993. An empirical study of hybrid inspection systems and allocation of inspection functions. International Journal of Human Factors in Manufacturing 351–367.

Jian, J., Bisantz, A.M., Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics 4 (1), 53–71.

Jiang, X., Bingham, J., Master, R., Gramopadhye, A.K., Melloy, B., 2002. A visual inspection simulator for hybrid environments. International Journal of Industrial Engineering: Theory, Applications and Practice 9 (2), 123–132.

Jiang, X., Gramopadhye, A.K., Melloy, B., Grimes, L., 2003. Evaluation of best system performance: human, automated and hybrid inspection systems. International Journal of Human Factors in Manufacturing 13 (2), 137–152.

Jiang, X., Gramopadhye, A.K., Melloy, B., 2004. Theoretical issues in the design of visual inspection systems. Theoretical Issues of Ergonomics Science 5 (3), 232–274.

Lee, J., Moray, N., 1992. Trust, control strategies and allocation of function in human-machine systems. Ergonomics 35, 1243–1270.

Master, R., Bingham, J., Jiang, X., Gramopadhye, A.K., Melloy, B., 2000. A questionnaire for measuring trust in hybrid inspection systems. Proceedings of the Industrial Engineering Research Conference, Dallas, TX.

Moray, N., Rodriguez, D., 2000. Levels of automation in process control. Proceedings of the IEA 2000/HFES 2000 Congress.

Morris, N., Rouse, W., Ward, S., 1988. Studies of dynamic task allocation on an aerial search environment. IEEE Transactions on Systems, Man and Cybernetics 18, 376–389.

Muir, B.M., 1994. Trust in automation. Part 2. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics 37, 1905–1922.

Muir, B.M., Moray, N., 1996. Trust in automation. Part 1. Experimental studies of trust and human intervention in a process control simulation. Ergonomics 39, 429–460.

Parasuraman, R., 2000. Designing automation for human use: empirical studies and quantitative models. Ergonomics 43, 931–951.

Rempel, J.K., Holmes, J.G., Zanna, M.P., 1985. Trust in close relationship. Journal of Personality and Social Psychology 49, 95–112.

Sanders, M.S., McCormick, E.J., 1993. Human Factors in Engineering and Design. McGraw-Hill, New York, pp. 120–121.

Sharit, J., Elhence, S., 1987. Technical and human aspects of computer-aided manufacturing. In: Salvendy, G. (Ed.), Handbook of Human Factors. Wiley, New York, pp. 1243–1268.

Sheridan, T.B., 1980. Computer control and human alienation. Technology Review 61–73.

Sheridan, T.B., Hennessy, R.T. (Eds.), 1984. Research and Modeling of Supervisory Control Behavior. National Academy Press, Washington, DC.

Sheridan, T.B., Fischhoff, B., Posner, M., Pew, R.W., 1983a. Supervisory control system. Research Needs for Human Factors. National Academy Press, Washington, DC.

Sheridan, T.B., Vamos, T., Aida, S., 1983b. Adapting automation to man, culture and society. Automatica 19 (6), 605–612.

Sinclair, M.A., 1984. Ergonomics of quality control. Workshop Document, International Conference on Occupational Ergonomics.

Sinclair, M.A., 1993. Human factors, design for manufacturability and the computer-integrated manufacturing enterprise. In: Helander, M., Nagamachi, M. (Eds.), Design for Manufacturability. Taylor & Francis, London, pp. 127–146.

Singh, I.L., Molloy, R., Parasuraman, R., 1993. Automation-induced ''complacency'': development of the complacency-potential rating scale. The International Journal of Aviation Psychology 3, 111–122.

Swets, J.A. (Ed.), 1964. Signal Detection Theory and Recognition by Human Observers: Contemporary Readings. Wiley, New York.

Thapa, V.B., Gramopadhye, A.K., Melloy, B.J., 1996. Evaluation of different training strategies to improve decision-making performance in inspection. The International Journal of Human Factors in Manufacturing 6 (3), 243–261.